Summary on
Smartphone – based Acoustic Indoor Space Mapping

Lohitanvita Rompicharla

RUID: 211009876

Department of Electrical and Computer Engineering

Objective

Indoor space mapping is one of the most important structural information for various modern applications today. It allows us to position objects, people, enhances indoor localization, navigation, offering valuable location-based services. There are currently many approaches like vision-based, LiDAR, multi-depth sensors, cameras, crowdsourcing, acoustic-based etc., for mapping indoor space. But all of them have their own limitations whether it may be cost, availability, privacy issues or accuracy. But this novel project "Smartphone-based Acoustic Indoor Space Mapping" has overcome many of these limitations providing efficient results.

Introduction

This research paper on "Smartphone-based Acoustic Indoor Space Mapping (SAMS)" is an amalgamation of many existing indoor mapping approaches. The authors have closely analyzed all the existing techniques and their drawbacks like

(i) LiDAR, which is costly and sensitive to transparent materials
(ii) Crowdsourcing, which is less expensive but slow, unsecure and disincentive
(iii) State-of-the-art "BatMapper", which is similar to this paper but is less accurate, needs multiple walks, and cannot detect contour for arbitrary shapes

However, this latest method uses a smartphone (Samsung Galaxy S7) with its inbuilt speaker, microphone, Frequency Modulated Continuous Wave (FMCW) technology for audio chirp, customized IMU, and MATLAB for analyzing and constructing the final contour. This infrastructure-free, calibration-free, acoustic-based technology supports distance measurement of single wall to multiple walls, straight and curved surfaces, accounts for clutters and corners in various environmental conditions. The accuracy of this advanced technology ranges from median errors of 1.5cm for a single wall to 6cm for multiple walls and 90th percentile of 1m in case of longer trajectory and presence of clutter, with just a single walk around the room. Also, it is inexpensive, fast, and easy to use since it does not require any external hardware. These astonishing features of SAMS enables it to have wide range of applications such as navigation for blind people, enriching AR/VR applications, indoor construction, wireless received signal strength prediction etc.

Working

The working of SAMS system is very simple yet efficient. Here, a user walks around an indoor space carrying a smartphone in hand that emits FMCW based audio chirps (acoustic signal) and records the received audio signal from different surfaces like walls, ceiling, floor, and objects along with IMU sensor data simultaneously. This measured distances from the nearby surfaces and objects are synchronized with user movement and trajectory to get the accurate data which is then processed in MATLAB by applying certain geometric constraints to construct an indoor space contour. There are mainly three modules to be considered in the SAMS system for full contour construction, namely

i) distance estimation

ii) dead reckoning module
iii) contour construction

For distance estimation, a FMCW chirp is transmitted periodically from the speaker and reflected chirps are recorded by the microphone simultaneously. We get a mixed signal given as $S_m(t) = \alpha cos(2\pi f_{min}\frac{R}{v_s} + 2\pi Bt\frac{R}{v_s T} - \pi B\frac{R^2}{v_s^2 T})$ where B is bandwidth, T is sweep time, which is correlatedly synchronized and passed through Hanning window, then by applying FFT on this multiplicative mixture of generated and received signal we obtain FMCW profile to estimate the distance. The key parameters effecting FMCW performance are a) bandwidth (resolution), b) chirp duration (range), c) peak selection, and d) volume. For this research experiment, bandwidth of 10KHz (11 to 21KHz) and chirp duration of 30ms has been selected because B=10KHz can easily differentiate reflectors that are separated by 10cm leading to lower merging of peaks and cd=30ms gives range of approximately 5.19m one way which is compatible with the user movement. Then comes peak selection, which is affected by factors like multi-path propagation, noise, phone orientation, user speed etc. Therefore, using Hanning window $H[n] = 0.5 * \left(1 - \cos\left(\frac{2*\pi*n}{N-1}\right)\right)$ improves the peak-to-side-lobe ratio and helps distinguish the main peak. To select the correct maximum peak, a peak selection algorithm has been described where latest peaks are grouped with previous peaks which are within a threshold T = 10cm, and at the end, the group with maximum number of peaks assigned to it is chosen for contour construction. To decide speaker volume, distances are measured from walls about 0.5 to 4m away in different indoor locations and volume level 5(30% of max) gave the best result with maximum Signal-to-Noise Ratio. A critical observation in constructing the contour is that irrespective of phone direction, SAMS measures perpendicular distance to the wall for both vertical & horizontal angles showing robustness of the system. Later, an algorithm is described for corner detection where the merged peaks within the range of 50cm of maximum peak are detected and then eliminated for contour construction.

The dead reckoning module estimates the user's trajectory i.e., step size, step count and heading direction using smartphone IMU (Accelerometer, Gyroscope, Magnetometer). Here, the vertical component of acceleration change that passed through frequency-based low-pass filter is calculated. This filtered data is then analyzed using peak recognition algorithm with a sliding window to detect peaks. Assuming that the user moves at a constant speed, step-size is calculated as $0.98 * \sqrt{A_{max} - A_{min}}$ where A is linear acceleration in a step duration. Later, by multiplying sensor values with rotation matrix, we identify the turns taken by the user and the heading direction is extracted.

Final step is contour construction. This involves wall association, wall co-ordinate estimation and outlier removal as crucial factors. The wall association algorithm uses each measurement (x_t, y_t, r_t) from dead-reckoning and FMCW modules to associate to a wall in the room. Then the gradients [$m_2 = \tan(\Theta + \tan^{-1}(m_1))$] are calculated, where Θ is the phone orientation ($\Theta = \sin^{-1}(r_{t+1} - r_t / (y_{t+1} - y_t)^2 + (x_{t+1} - x_t)^2$). Then with similar gradients grouped we get whether it is a straight wall, curved wall, or an intersection, which helps in getting the co-ordinates of the wall. Thereafter, using linear regression and Cook's distance ($D_i > 4/n$), the outliers are identified and removed. To reconstruct the actual contour from generated coordinates, root-mean-squared-error (RMSE) is performed i.e., if $RMSE_{linearregression}$ - $RMSE_{splineinterpolation} \geq T$ with threshold T= 8cm, then it is a curved wall otherwise a straight wall. The point of intersection between different consecutive walls gives the estimated coordinates of corners. With all the equations of FMCW profile, wall and corner coordinates handy, a room contour can be plotted well.

Setup

The whole SAMS system was tested in various setups and environmental conditions. Initially all the testing setup was made with wooden structures made as single, multi-walls and curved walls with corners and clutters. The experiment was also conducted in an anechoic chamber with laser-based systems to get the ground truth. Also, Samsung Galaxy S4 was used to reduce the user profiling effect as its speaker is at the rear side and microphone at two ends. In all these scenarios the systems showed accurate results with a median error of 6cm.

Results

Taking various external factors such as HVAC system, people, loud noise into account, the results have been established as follows: SAMS systems is robust to ambient sound, the effect of Doppler shift is negligible, and for clutters more than 60cm width that can block the direct path gives an error of 0.2cm for 80% of values.

Comparison

Now, comparing the state-of-the-art technology BatMapper with SAMS system, the main differences spotted are: i) it requires extensive training to gather the parameters for algorithm unlike SAMS which is calibration-free, ii) it uses echo-based distance estimation which has higher error compared to SAMS, iii) cannot construct contour for arbitrary shapes, iv) requires multiple trials to build a map of single space which is contradictory to SAMS that can construct whole contour for arbitrary shapes, closed room with numbers of walls etc., in a single walk.

Conclusion

We can conclude that the "Smartphone-based Acoustic Mapping System" gives promising results with high accuracy and less errors with all the external and internal factors considered. We can obtain the real time contour construction of any indoor space i.e., from closed room to corridors with corners and clutters as well. The paper also presented a use-case that represents that SAMS can be used to predict the Received Signal Strength (RSS) (within 1.5-2dB error) at any given location which can be applied to wireless optimization tasks like AP replacement, AP selection and rate adaptation.

Future Scope

Given all the advantages, disadvantages and working of SAMS system, we can conclude that the system has a wide number of applications in the present world and further it can be improved and tested for getting more accurate contour construction even in the presence of large clutters and used for specific applications like indoor localization, navigation, and AR/VR enrichment.